Nucleic Acids Research, 1994, Vol. 22, No. 9 1593–1599

Codon cassette mutagenesis: a general method to insert or replace individual codons by using universal mutagenic cassettes

Deena M.Kegler-Ebo, Catherine M.Docktor and Daniel DiMaio*
Department of Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510, USA

Received February 10, 1994; Revised and Accepted March 28, 1994

ABSTRACT

We describe codon cassette mutagenesis, a simple method of mutagenesis that uses universal mutagenic cassettes to deposit single codons at specific sites in double-stranded DNA. A target molecule is first constructed that contains a blunt, double-strand break at the site targeted for mutagenesis. A double-stranded mutagenic codon cassette is then inserted at the target site. Each mutagenic codon cassette contains a three base pair direct terminal repeat and two head-to-head recognition sequences for the restriction endonuclease *SapI*, an enzyme that cleaves outside of its recognition sequence. The intermediate molecule containing the mutagenic cassette is then digested with *SapI*, thereby removing most of the mutagenic cassette, leaving only a three base cohesive overhang that is ligated to generate the final insertion or substitution mutation. A general method for constructing blunt-end target molecules suitable for this approach is also described. Because the mutagenic cassette is excised during this procedure and alters the target only by introducing the desired mutation, the same cassette can be used to introduce a particular codon at all target sites. Each cassette can deposit two different codons, depending on the orientation in which it is inserted into the target molecule. Therefore, a series of eleven cassettes is sufficient to insert all possible amino acids at any constructed target site. Thus codon cassettes are 'off-the-shelf' reagents, and this methodology should be a particularly useful and inexpensive approach for subjecting multiple different positions in a protein sequence to saturation mutagenesis.

INTRODUCTION

Site-directed mutagenesis has proven to be a powerful tool for analyzing protein structure and function. A variety of methods have been developed for the introduction of specific amino acid changes at predetermined sites in proteins (1–3). In some cases, it is desirable to introduce multiple different substitutions at a particular position or at several positions in a protein and to determine the consequences of these changes on protein function. This type of saturation mutagenesis has been used, for example, to determine the amino acid requirements for activation of the Harvey-*ras* oncogene by substitutions at codon 12 or for activation of the *neu* oncogene by substitutions at codon 664 within the transmembrane domain of $p185^{neu}$ (4–6).

Saturation mutagenesis is often carried out by using oligonucleotide-directed mutagenesis (2,3,7). In this approach, a template molecule is usually prepared by cloning the target gene or a fragment thereof into a phage M13-based single-stranded cloning vector. Mutagenesis is then performed by hybridization of a mutagenic oligodeoxyribonucleotide, enzymatic extension from this oligonucleotide primer, and ligation prior to transformation of bacteria, followed by screening to identify the mutants. By using a panel of mutagenic oligonucleotides with different mismatches at the targeted codon, a position can be readily saturated with mutations. However, the cost of oligonucleotide synthesis can rapidly become prohibitive, especially if multiple positions are to be mutagenized, since a different set of mutagenic primers must be synthesized for each position. Alternatively, degenerate oligonucleotide mixtures with the potential to generate several different mismatches can be used as mutagenic primers (8,9). Although this approach reduces the cost of oligonucleotide synthesis, an extensive screening effort is required to identify all possible substitutions, and the yield of new mutations often drops dramatically as the set of mutations isolated nears completion. Furthermore, a different degenerate mutagenic oligonucleotide must be synthesized to mutagenize each targeted region of the gene. PCR-based approaches to introduce specific mutations into DNA have also been described, but in general specific mutagenic oligonucleotide primers must be synthesized for each mutation (10,11).

Cassette mutagenesis methods are also useful, particularly in cases where a targeted region is contained on a small DNA fragment flanked by restriction endonuclease cleavage sites that do not cut elsewhere in the target plasmid (12–14). Cleavage with these enzymes can be used to excise a small segment of DNA which is then replaced with a single- or double-stranded

*To whom correspondence should be addressed

DNA segment containing a defined mutation or a mixture of mutations. The requirement for nearby unique sites (or sequences that can be converted into unique sites by mutagenesis) has been overcome by the use of *Bsp*MI cassettes (15). In this approach, the region to be mutagenized is replaced with a DNA segment containing two *Bsp*MI recognition sequences arranged in opposite orientations. Since *Bsp*MI cleaves outside its recognition sequence, *Bsp*MI cleavage removes a segment of DNA, leaving an acceptor molecule into which an oligonucleotide can be inserted to restore the normal sequence, except for the mutagenized position. Again, the cost of oligonucleotide synthesis or screening can be prohibitive for cassette mutagenesis, particularly if multiple regions of a gene are subjected to saturation mutagenesis, since a different set of oligonucleotide cassettes harboring the codon changes must be produced for each mutagenized region.

In addition to these methods which target specific codons, a variety of chemical and enzymatic methods have also been described to subject a short sequence of DNA to intense mutagenesis (e.g. 16). Such methods frequently entail complex *in vitro* manipulations or extensive screening efforts to identify the desired mutations, and often multiple mutations are introduced into a single molecule.

We have developed codon cassette mutagenesis, a new method that utilizes a set of eleven universal mutagenic cassettes to insert or replace individual codons at blunt-end target sites in DNA molecules. The desired codon change is the only remnant of the cassette left in the target molecule, so the same set of cassettes can be used to introduce mutations at all blunt-end targets. We also describe a related approach to construct a blunt-end at any position in DNA, which thereby allows codon cassette mutagenesis to be employed at any position in a gene, regardless of the DNA sequence context. Thus, a single series of cassettes can be used to introduce codons encoding all possible amino acids at targets that can be constructed at all positions in any gene.

MATERIALS AND METHODS

Plasmid DNA construction

pUC19 plasmid DNA was digested to completion with *Tfi*I, and the 2.5 kb fragment lacking the unique *Sap*I recognition sequence at nucleotide 683 was recircularized to generate pUC19-Sap. A double-stranded oligonucleotide cassette containing *Bam*HI- and *Pst*I-compatible single-strand ends was ligated to the large *Bam*HI−*Pst*I fragment of pUC19-Sap to generate pUC19-Sap-T, as shown in Figure 4.

Oligodeoxyribonucleotides

Codon cassette oligonucleotides were the generous gift of New England Biolabs, Inc., Beverly, MA. They all have the basic structure shown in Fig. 1B, differing only in the direct terminal repeats. Other oligonucleotides were obtained from the Keck Biotechnology Laboratory at Yale University School of Medicine, New Haven, CT. To form double-stranded cassettes, complementary oligonucleotides (3.5 ng/μl each in 1.0 mM Tris, pH 7.5, 0.05 mM $MgCl_2$) were annealed at a 1:1 ratio by warming at 65°C for 1−2 min, followed by slow cooling to 30°C.

Ligations

Approximately 250 ng of linearized pUC19-Sap or filled-in pUC19-Sap-T were ligated to double-stranded cassettes (28 ng) by using 400 units of T4 DNA ligase in 10 μl total volume. Blunt-end ligations were incubated at 16°C overnight, and cohesive-end ligations were incubated at room temperature for 1 h. In some cases, blunt-end ligation reactions were digested with a restriction endonuclease to enrich for molecules that had lost a restriction cleavage site.

*Sap*I digestion

Approximately 200 ng of miniprep DNA (17) was incubated with 1 unit of *Sap*I (New England Biolabs, Inc.) in a 20 μl reaction for 1.5 h at 37°C. The reaction products were diluted to approximately 5−10 μg/ml and circularized by ligation.

General procedures

Standard procedures were used to transform DH5α *E.coli* and to screen for β-galactosidase expression (18). All restriction endonucleases and Klenow DNA polymerase were purchased from New England Biolabs, Inc., Beverly, MA, except for *Sap*I endonuclease, which was a gift from the company, and all were used according to manufacturer's instructions. DNA sequences were determined by using the chain termination method (19), the Sequenase Kit version 2.0 (United States Biochemical Corporation), and New England Biolabs primers # 1233 and # 1212 for Figures 3 and 5, respectively.

RESULTS

Strategy

Type IIs restriction endonucleases cleave at sites adjacent to their recognition sequence. We reasoned that it should be possible to temporarily insert an element bearing recognition sequences for such an enzyme in the middle of a gene and then to remove these recognition sequences by digestion with the enzyme, leaving behind a desired mutation. Two features of *Sap*I allowed it to be used in this fashion to introduce single codons into DNA (Figure 1A). First, *Sap*I cleaves outside of its recognition sequence, and second, the sites of cleavage are one and four bases from the end of the recognition sequence, thereby leaving a three base overhang. Because the bases in the overhang are not part of the *Sap*I recognition sequence, any nucleotide sequence can comprise the overhang. Based on these properties of *Sap*I, we designed codon cassette mutagenesis, a simple procedure that

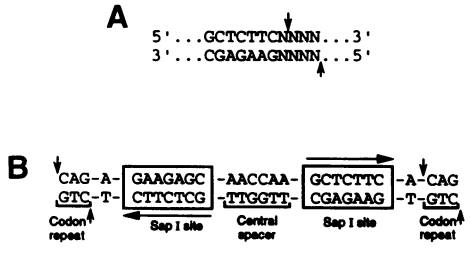

Figure 1. A. *Sap*I recognition sequences and cleavage sites. The *Sap*I recognition sequence in double-stranded DNA is shown, with the sites of strand scission indicated by the vertical arrows. N indicates that a position can be occupied by any base. **B.** Structure of a universal codon cassette. See text for description of its component parts. The *Sap*I recognition sequences are shown in the boxes, and the horizontal arrows point toward the cleavage sites. Sites of *Sap*I cleavage are shown by vertical arrows. The direct terminal codon repeat and central spacer between the two *Sap*I recognition sequences are indicated. The cassette shown introduces CAG (glutamine) and CUG (leucine) codons.

utilizes universal mutagenic cassettes to insert or replace single codons in cloned genes. First, a blunt-end at the site selected for mutagenesis is constructed in a target molecule devoid of *Sap*I recognition sequences. A mutagenic codon cassette is then inserted into the target molecule at the blunt-end site and subsequently removed, leaving the desired mutation in the target DNA. A mutagenic codon cassette can function in this manner because it contains two *Sap*I recognition sequences oriented so that *Sap*I cleavage occurs near the ends of the cassette (Figure 1B). As a consequence of this arrangement, digestion with *Sap*I removes all of the cassette from the target molecule except for complementary three base overhangs which are subsequently ligated to generate the single codon mutation. Thus, the only permanent sequence alteration caused by this procedure is the desired mutation, and one mutagenic cassette can be used to introduce the same codon at any blunt-end target site. Codon insertion using this approach is described first, then a modification of the procedure is presented which is a general method to construct target molecules suitable for codon replacement.

Design of the mutagenic codon cassette

The basic design of mutagenic codon cassettes is shown in Figure 1B. Double-stranded codon cassettes are fully base-paired and are generated by annealing two complementary oligodeoxyribonucleotides. Because codon cassettes have blunt-ends, they can be inserted into any blunt-end target molecule. Each codon cassette contains a three base pair direct terminal repeat corresponding to the codon that is eventually introduced. Each cassette also contains two head-to-head *Sap*I recognition sequences positioned so that, once the cassette is inserted into a target molecule, *Sap*I cleaves on each side of the three base pair terminal repeats, thereby generating a three base cohesive overhang attached to each end of the target molecule. In order for cleavage to occur at the proper position, an extra base pair is inserted between the end of each *Sap*I recognition sequence and each three base pair repeat. To minimize the chance that one of the constituent single-stranded oligonucleotides will form a homoduplex or form a hairpin during the annealing step, a non-palindromic central spacer is also included between the two *Sap*I recognition sequences. The oligonucleotides are not phosphorylated prior to ligation to the target plasmid so that the insertion of a single mutagenic cassette occurs.

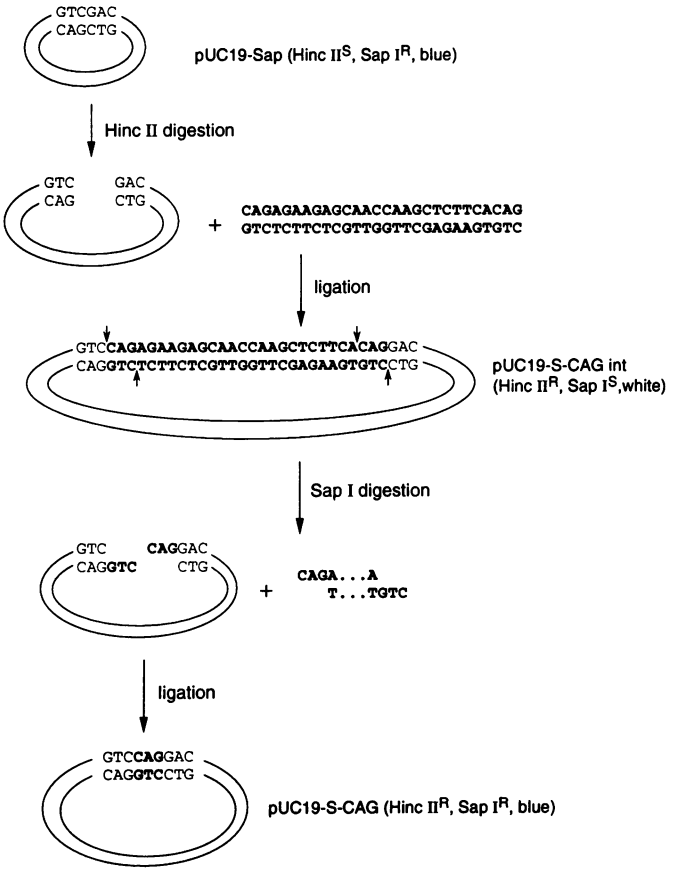

Figure 2. Insertion of a CAG codon at the *Hinc*II site of pUC19-Sap. The universal codon cassette is shown in boldface, as is the CAG triplet introduced by cassette ligation and resolution. Only one of two possible orientations of the inserted cassette is shown. The resistance (R) or sensitivity (S) to pertinent restriction endonucleases is shown in parentheses, as is the color of colonies induced by the various plasmids upon plating in the presence of X-gal. Sites of *Sap*I cleavage are shown by the short vertical arrows.

Codon insertion

We first tested this approach to insert single codons into the unique *Hinc*II site of pUC19-Sap, a derivative of pUC19 that contains no *Sap*I cleavage sites (Figure 2). *Hinc*II cleavage generates blunt-ends so no further manipulations were required to construct the target. A mutagenic codon cassette with a CAG terminal repeat was ligated into *Hinc*II-digested pUC19-Sap, resulting in the generation of pUC19-S-CAGint if the cassette is inserted in the orientation shown in Figure 2 or of pUC19-S-CTGint if it is inserted in the opposite orientation. Because the *Hinc*II site is located in the lacZ α gene and insertion of the 22 base pair mutagenic codon cassette results in a frameshift mutation, the insertion and resolution of the cassette was readily assessed by colony color following the plating of transformed

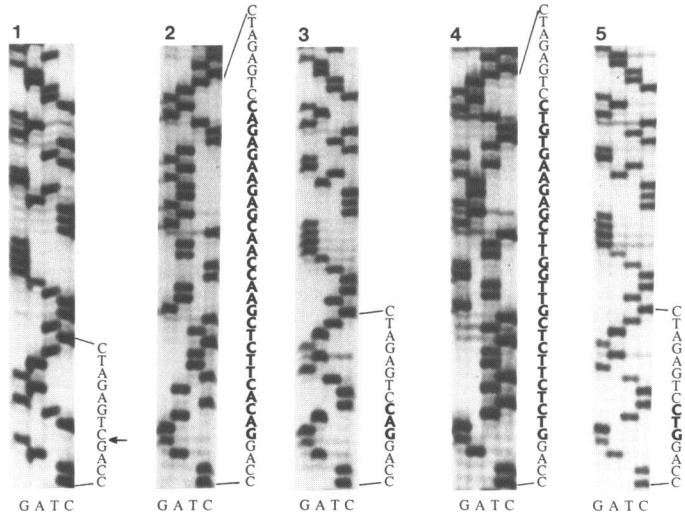

Figure 3. Sequences of codons inserted by using codon cassette mutagenesis. Sequences of plasmid DNA were determined as described in Materials and Methods. Reactions were loaded in the order GATC, and the sequence determined is shown at the right of each autoradiogram. Sequences should be read from top-to-bottom to correspond to Figure 2. Site of *Hinc*II cleavage is indicated by the arrow in panel A. Nucleotides inserted are shown in boldface. The following templates were used: (1) pUC19-Sap; (2) pUC19-S-CAGint; (3) pUC19-S-CAG; (4) pUC19-S-CTGint; (5) pUC19-S-CTG.

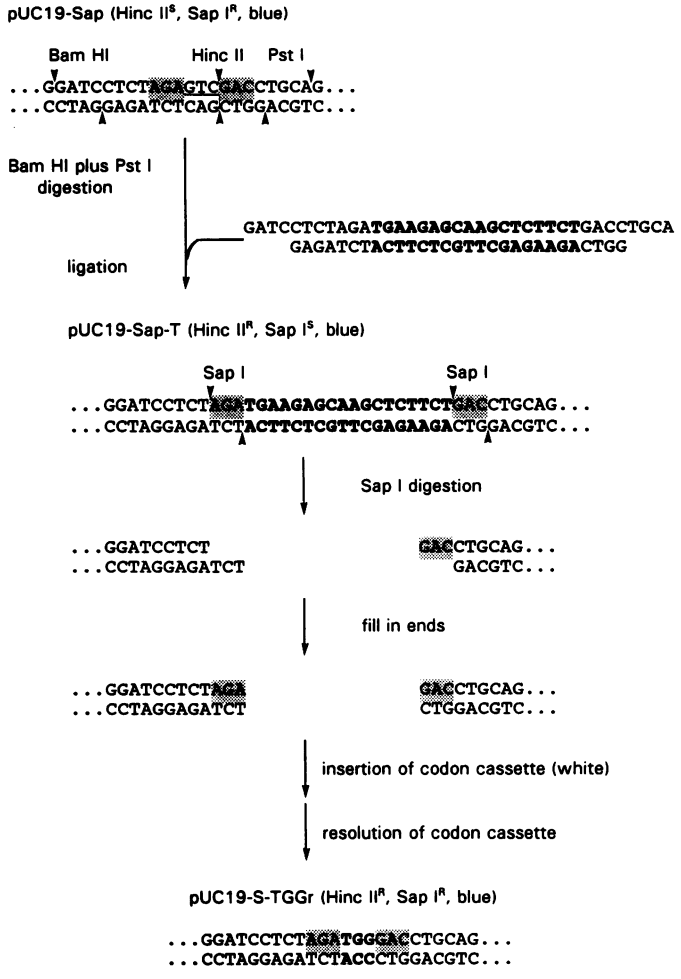

Figure 4. Replacement of codons by using codon cassette mutagenesis. See text for a more detailed description of this procedure. The GTC triplet targeted for replacement is underlined at the top of the figure, and the triplets flanking this triplet are shown with a stippled background. The cassette with the *Pst*I and *Bam*HI ends was ligated into *Bam*HI- and *Pst*I-digested pUC19-Sap to generate pUC19-Sap-T, in which the GTC triplet is replaced with head-to-head *Sap*I recognition sequences. pUC19-Sap-T was then digested with *Sap*I and the resulting overhanging ends were filled in by using the Klenow fragment of DNA polymerase. A universal TGG/CCA (tryptophan/proline) mutagenic codon cassette was then inserted and resolved as outlined in Figure 2. Inserted sequences that are not identical to wild type pUC19 are shown in boldface. The resistance (R) or sensitivity (S) to pertinent restriction endonucleases is shown in parentheses, as is the color of colonies induced by the various plasmids upon plating in the presence of X-gal. The sites of cleavage by various restriction endonucleases are shown by vertical arrowheads.

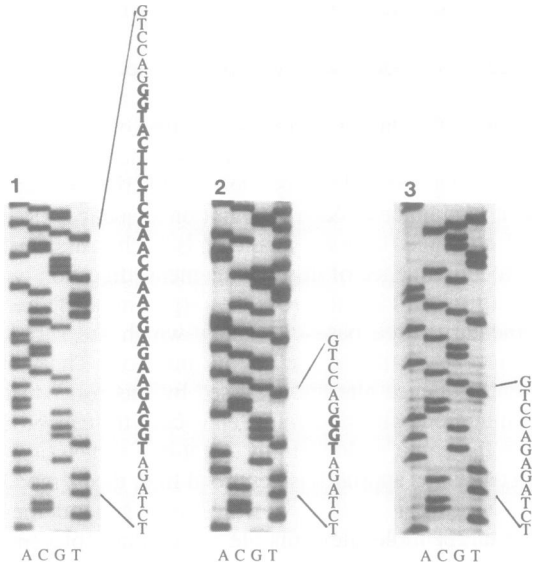

Figure 5. Sequences of codon replacement or deletion mutations constructed by using codon cassette mutagenesis. Reactions were loaded in the order ACGT and should be read from bottom-to-top to correspond to Figure 4. The following templates were used: (1) pUC19-ST-TGGint, containing a TGG/CCA mutagenic cassette in place of targeted codon; (2) pUC19-ST-TGGr, containing a TGG replacement; (3) pUC19-ST-dl, containing a deletion of the targeted codon.

bacteria on agar containing 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and isopropylthiogalactoside (IPTG). Colonies containing a plasmid with an inserted cassette were identified by their white appearance on indicator plates, and restriction analysis of plasmid DNA confirmed the loss of the *Hinc*II site and acquisition of *Sap*I sites. Isolates with the cassette inserted in each orientation at the targeted site were identified by sequence analysis of plasmid miniprep DNA. DNA was then digested to completion with *Sap*I, diluted, and recircularized via the three base pair cohesive ends. Plasmids from which the cassette had been successfully excised induced the formation of blue colonies on indicator plates. The removal of the cassette was confirmed by resistance of plasmid DNA to *Sap*I digestion, and the precise insertion of the desired codon was confirmed by DNA sequencing. Figure 3 shows the DNA sequences of the starting plasmid with the wild type *Hinc*II site (panel 1), the intermediate plasmids with the cassette inserted in either orientation (panels 2 and 4), and the final plasmids in which the cassette is resolved to yield each of the single codon insertions (panels 3 and 5). Thus, this method can be used as predicted to insert three base pairs into DNA. In addition to the successful use of the CAG/CTG (glutamine/leucine) cassette as shown in Figure 3, we have obtained single codon insertions by using ATG/CAT (methionine/histidine) and TGG/CCA (tryptophan/proline) cassettes of the same design. We have also constructed single codon insertions with three other cassettes with a related design.

Although we used indicator plates to identify plasmids with the appropriate sequence changes in the experiments described above, this was not necessary because most colonies displayed the desired phenotype after each step of the procedure (i.e. white after ligation of the cassette and blue after *Sap*I digestion and recircularization). Furthermore, colony hybridization to an oligonucleotide probe complementary to the central portion of the mutagenic cassette documented that >95% of white colonies did, in fact, contain a cassette (data not shown). Therefore, random screening of colonies is a feasible approach to identify the correct intermediates and mutants constructed by using this method. In addition, it is possible to use restriction endonuclease digestion to enrich for molecules containing an inserted cassette because it destroys the original blunt-end recognition sequence site (in this case for *Hinc*II) and also for molecules from which the cassette has been excised because the *Sap*I sites are removed. In practice, we have readily obtained codon insertions by using

this method without resorting to color screening, other phenotypic selection, or enzymatic enrichment.

Target molecule construction and codon replacement

For many applications, it is desirable to replace one codon with another or with a series of different codons. To replace codons using codon cassette mutagenesis, the target plasmid is first manipulated so that blunt-ends flank the position selected for replacement, i.e. one blunt-end is immediately downstream of the codon that is 5′ to the targeted position and the other is immediately upstream of the codon that is 3′ to this position. The set of universal codon cassettes is then used with this target molecule to introduce the codon replacement. In this section, we describe a general method to generate such targets and obtain substitution mutations. This method entails first constructing a DNA molecule in which the codon targeted for substitution is replaced by a DNA segment containing two *Sap*I recognition sequences. These recognition sequences are positioned so that the sites of *Sap*I strand scission bracket each of the two codons that originally flanked the targeted codon. The desired blunt-ended target molecule is then generated by cleaving with *Sap*I and filling in the resulting single-strand ends of the target molecule with DNA polymerase. The resulting target molecule thus contains blunt-ends flanking a deletion of the codon targeted for replacement. Mutagenic codon cassettes are then inserted into this target site and resolved by *Sap*I digestion and religation, resulting in the conversion of the original three base pair deletion into a codon replacement.

This approach was successfully used as shown in Figure 4 to generate substitutions of the GUC codon at the 5′ side of the *Hin*cII cleavage site of pUC19-Sap (shown as the underlined GTC triplet in the figure). pUC19-Sap-T was first constructed by using standard cassette mutagenesis to replace the segment of the pUC19-Sap polylinker between the *Bam*HI and *Pst*I cleavage sites with a segment of DNA that reconstituted the wild type pUC19 sequences, except for head-to-head *Sap*I recognition sequences in place of the GUC codon targeted for substitution. A spacer nucleotide was included adjacent to each *Sap*I recognition sequence so that the sites of *Sap*I cleavage occur on either side of each of the two codons that originally flanked the GUC codon. Insertion of this new sequence and *Sap*I cleavage thus result in the deletion of the targeted codon and the generation of single-stranded ends corresponding to the two flanking codons. These ends were then filled-in by using the Klenow fragment of DNA polymerase and all four deoxyribonucleoside triphosphates, thereby generating the deleted target molecule with blunt-ends at the appropriate positions as shown in Figure 4. Codons were then inserted into this target by using a universal mutagenic codon cassette as described above, resulting in the replacement of the original targeted codon with the desired ones. In the example shown, a TGG/CCA (tryptophan/proline) cassette was used to generate pUC19-ST-TGGint (Figure 5, panel 1), and the final products with the desired single codon replacements were readily obtained (Figure 5, panel 2). Despite the inclusion of the additional enzymatic step required to generate blunt-ends, color screening on indicator plates and direct sequence analysis indicated that this cassette was correctly inserted in >60% of the colonies examined and that it was correctly resolved in approximately 80% of the cases. It should also be noted that deletion of the targeted GUC codon can be easily achieved by recircularization of the filled-in, blunt-end target plasmid without insertion of the codon cassette (Figure 5, panel 3).

Table 1A.

Amino acid	Codon	Restriction endonuclease	Recognition sequence
Alanine	GCC	*Nae*I	**GCC/GGC**
Arginine	AGG	*Stu*I	**AGG/CCT**
Asparagine	AAU	*Ssp*I	**AAT/ATT**
Aspartic acid	GAU	*Eco*RV	**GAT/ATC**
Cysteine	UGC	*Fsp*I	**TGC/GCA**
Glutamic acid	GAG	*Bsr*BI	**GAG/CGG**
	GAG	*Ecl*136II	**GAG/CTC**
Glutamine	CAG	*Pvu*II	**CAG/CTG**
Glycine	GGC	*Ehe*I	**GGC/GCC**
Histidine	CAC	*Pml*I	**CAC/GTG**
Phenylalanine	UUU	*Dra*I	**TTT/AAA**
Proline	CCC	*Sma*I	**CCC/GGG**
	CCG	*Msp*A1I	**CCG/CGG**
	CCG	*Bsr*BI	**CCG/CTG**
Serine	UCG	*Nru*I	**TCG/CGA**
	AGC	*Eco*47III	**AGC/GCT**
	AGU	*Sca*I	**AGT/ACT**
Tryptophan	UGG	*Msc*I	**TGG/CCA**
Tyrosine	UAC	*Sna*BI	**TAC/GTA**
Valine	GUU	*Hpa*I	**GTT/AAC**
	GUA	*Bst*1107	**GTA/TAC**
	GUC	*Hin*cII	**GTC/GAC**

The table shows codons (and their encoded amino acids) that immediately precede a blunt-end generated by cleavage by the indicated restriction endonuclease. The nucleotides in the recognition sequence that correspond to the codon are shown in boldface. Note that codons for isoleucine, leucine, lysine, methionine, and threonine are not accommodated by the available restriction endonucleases.

Table 1B.

Amino acid	Codon	Restriction endonuclease	Recognition sequence
Alanine	GCC	*Ehe*I	**GGC/GCC**
	GCA	*Fsp*I	**TGC/GCA**
	GCU	*Eco*47III	**AGC/GCT**
Arginine	CGA	*Nru*I	**TCG/CGA**
	CGG	*Bsr*BI	**GAG/CGG**
	CCG	*Msp*A1I	**CCG/CGG**
Asparagine	AAC	*Hpa*I	**GTT/AAC**
Aspartic acid	GAC	*Hin*cII	**GTC/GAC**
Glycine	GGC	*Nae*I	**GCC/GGC**
	GGG	*Sma*I	**CCC/GGG**
Isoleucine	AUC	*Eco*RV	**GAT/ATC**
	AUU	*Ssp*I	**AAT/ATT**
Leucine	CUG	*Pvu*II	**CAG/CTG**
	CUC	*Ecl*136II	**GAG/CTC**
	CUC	*Bsr*BI	**CCG/CTC**
Lysine	AAA	*Dra*I	**TTT/AAA**
Proline	CCA	*Msc*I	**TGG/CCA**
	CCU	*Stu*I	**AGG/CCT**
Threonine	ACU	*Sca*I	**AGT/ACT**
Tyrosine	UAC	*Bst*1107	**GTA/TAC**
Valine	GUA	*Sna*BI	**TAC/GTA**
	GUG	*Pml*I	**CAC/GTG**

The table shows codons (and their encoded amino acids) that immediately follow a blunt end generated by cleavage by the indicated restriction endonuclease. The nucleotides in the recognition sequence that correspond to the codon are shown in boldface. Note that codons for cysteine, glutamic acid, glutamine, histidine, methionine, phenylalanine, serine, and tryptophan are not accommodated by the available restriction endonucleases

DISCUSSION

Codon cassette mutagenesis uses a set of universal oligodeoxyribonucleotide cassettes to deposit single codons into blunt-end target molecules. This is a general method of mutagenesis because

Table 2.

Mutagenic cassette[a]	Inserted amino acid[b]
ATG.....ATG	methionine
TAC.....TAC	histidine
TGG.....TGG	tryptophan
ACC.....ACC	proline
CAG.....CAG	glutamine
GTC.....GTC	leucine
GAC.....GAC	aspartic acid
CTG.....CTG	valine
AAC.....AAC	asparagine
TTG.....TTG	valine
TAT.....TAT	tyrosine
ATA.....ATA	isoleucine
GGC.....GGC	glycine
CCG.....CCG	alanine
AAA.....AAA	lysine
TTT.....TTT	phenylalanine
TTC.....TTC	phenylalanine
AAG.....AAG	glutamic acid
AGA.....AGA	arginine
TCT.....TCT	serine
ACA.....ACA	threonine
TGT.....TGT	cysteine

[a]Schematic version of a series of universal codon cassettes and their direct terminal repeats are shown. Each cassette has the general structure shown in Figure 1B, with the top strand written in the 5′ to 3′ direction, and the bottom strand in the 3′ to 5′ orientation.

[b]The amino acids inserted by the cassettes in the first column is shown. For each cassette, the amino acid on the top is introduced if the cassette is inserted in the orientation shown in the first column, and the amino acid on the bottom is introduced if the cassette is inserted in the opposite orientation.

suitable targets can be constructed at any position in a gene, and any amino acid in a protein can be replaced with all others. This is in contrast to other methods of codon insertion that in general require pre-existing restriction endonuclease cleavage sites or are compatible only with particular insertions (e.g. 20). The only requirement for using codon cassette mutagenesis is that the starting target plasmid contains no endogenous *Sap*I cleavage sites. This is not a difficult condition to meet since the seven base pair *Sap*I recognition sequence occurs relatively infrequently in DNA. Moreover, because of the degeneracy of the genetic code, it is possible to inactivate by point mutation a *Sap*I recognition sequence in any coding region without altering the encoded amino acids. The compelling advantage of codon cassette mutagenesis is that the same set of mutagenic codon cassettes can be used for codon insertion or replacement with all possible target molecules. Thus, once target molecules are constructed and the set of mutagenic codon cassettes is obtained, all possible single codon insertion or substitution mutations can be generated with no additional expense for oligonucleotide synthesis. The position of the blunt-ends in the target molecule determine whether the final mutation is a codon insertion or replacement. If the blunt-ends are located between two adjacent codons in the wild type sequence, a new codon will be inserted between these two codons (as shown in Figure 2), whereas if the blunt-ends are on both sides of the targeted codon, a codon replacement will occur (as shown in Figure 4).

The only mutation that must be custom-designed and constructed in codon cassette mutagenesis is the one that allows the generation of the blunt-ends at the site of codon insertion or replacement. The insertion of restriction endonuclease recognition sequences to allow the generation of these blunt-ends will frequently entail the use of standard cassette, PCR- or M13-based mutagenesis procedures. In favorable cases, restriction endonucleases with blunt-end cleavage sites can be used to generate blunt-ends at the appropriate positions. Table 1 lists codons that are located immediately 5′ or 3′ to the cleavage sites of restriction endonucleases that generate blunt-ends. However, as shown in Table 1, not all amino acids are accommodated by the available restriction endonucleases. Furthermore, an endonuclease that generates a desired blunt-end might cut elsewhere in the target plasmid, thereby necessitating various subcloning steps to remove extraneous sites and to reconstruct the final mutation into an expression vector for analysis. A more general method of target construction entails the use of *Sap*I recognition sequences as shown in Figure 4. The major advantage of using this approach, which is similar to the use of 'excision cassettes' as described by Palzkill and Botstein to generate blunt-ends for insertional mutagenesis (21), is that it allows construction of a blunt-end at any position without altering the amino acid sequences encoded by adjacent DNA. Furthermore, this approach imposes no limitations on the occurrence of restriction endonuclease cleavage sites other than *Sap*I sites elsewhere in the plasmid. *Sap*I recognition sequences can also be used in this fashion for other applications where it is desirable to construct a new blunt, double-strand break in a gene without altering the flanking amino acid sequence.

Once a target molecule is constructed, codon cassettes are inserted. Because a cassette can be inserted in two orientations, each cassette has the potential to introduce two different codons. Table 2 lists the direct terminal repeats of a series of 11 codon cassettes and the amino acids these cassettes insert. The use of codons rarely used in mammalian mRNAs was avoided in choosing these direct repeats. These cassettes in aggregate insert all 20 codons that encode amino acids. During a saturation mutagenesis procedure, each cassette is inserted in a separate ligation reaction, and all ligation reactions are processed in parallel to yield the final mutations. In our hands, we have found it practical to transform bacteria with each ligation reaction, isolate transformed colonies, and to use sequence analysis to identify isolates with cassettes inserted in either orientation. An isolate in each orientation is resolved separately to yield the final mutations. Alternatively, DNA can be subjected to cassette insertion followed by *Sap*I digestion without an intermediate transformation step, and the two mutations introduced by each cassette can be identified among the resulting progeny (data not shown). Although several simple enzymatic reactions must be carried out to obtain each pair of mutants using codon cassette mutagenesis, little screening is required to identify the desired mutations, since all the steps are very efficient and only two mutations are possible for each set of reactions. Furthermore, it should be possible to modify the design of codon cassettes, for example by including the lac operator between the *Sap*I sites, so that colonies that harbor a plasmid containing an inserted cassette can be directly identified on indicator plates.

Palzkill and Botstein have independently developed a related mutagenesis procedure, in which *Bsp*MI sites are temporarily inserted into target DNA and subsequently removed, resulting in the introduction of random nucleotide sequences and the replacement of 3–6 contiguous codons (21). Therefore, this method cannot be used for replacement or insertion of individual codons. In contrast, codon cassette mutagenesis was specifically designed to allow the rapid and economical generation of single amino acid substitution mutations. This approach is particularly well-suited for the saturation of individual positions in a protein with all possible amino acid substitutions. The ease with which such mutations are generated and identified suggests that codon cassette mutagenesis will facilitate the detailed structure/function analysis of interesting proteins.

ACKNOWLEDGEMENTS

We thank Richard Morgan and Ira Schildkraut of New England Biolabs, Inc. for gifts of essential reagents. We also thank Jan Zulkeski for assistance in preparing this manuscript. This work was supported by a grant from the National Cancer Institute (R37 CA37157).

REFERENCES

1. Botstein, D., and Shortle, D. (1985) *Science* **229**:1193–1201.
2. Itakura, K., Rossi, J.J., and Wallace, R.B. (1984) *Ann. Rev. Biochem.* **53**:323–356.
3. Smith, M. (1985) *Ann. Rev. Genet.* **19**:423–462.
4. McGrath, J.P., Capon, D.J., Goeddel, D.V., and Levinson, A.D. (1984) *Nature* **310**:644–649.
5. Seeburg, P.H., Colby, W.W., Capon, D.J., Goeddel, D.V., and Levinson, A.D. (1984) *Nature* **312**:71–75.
6. Bargmann, C.I., and Weinberg, R.A. (1988) *EMBO J.* **7**:2043–2052.
7. Zoller, M.J., and Smith, M. (1982) *Nucl. Acids Res.* **10**:6487–6500.
8. Hsiung, H., Inouye, S., West, J., Sturm, B., and Inouye, M. (1983) *Nucl. Acids. Res.* **11**:3227–3239.
9. Horwitz, B.H., and DiMaio, D. (1990) *Meth. Enzymol.* **185**:599–611.
10. Higuchi, R., Krummel, B. and Saiki, R. (1988) *Nucl. Acids Res.* **16**:7351–7367.
11. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989) *Gene* **77**:51–59.
12. Matteucci, M.D., and Heyneker, H.L. (1983) *Nucl. Acids Res.* **11**:3113–3121.
13. Wells, J.A., Vasser, M., and Powers, D.B. (1985) *Gene* **34**:315–323.
14. Derbyshire, K.M., Salvo, J.J., and Grindley, N.D. (1986) *Gene* **46**:145–152.
15. Stone, J.C., Vass, W.C., Willumsen, B.M., and Lowy, D.R. (1988) *Mol. Cell. Biol.* **8**:3565–3569.
16. Myers, R.M., Lerman, L.S., and Maniatis, T. (1985) *Science* **229**:242–246.
17. Goode, B.L., and Feinstein, S.C. (1992) *Biotechniques* **12**:373–374.
18. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
19. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
20. Boeke, J.D. (1981) *Mol. Gen. Genet.* **181**:288–291.
21. Palzkill, T., and Botstein, D. (1991) *Methods: A Companion to Methods in Enzymology* **3**:155–164.